



Comparative Legilinguistics

vol. 46/2021

DOI: <http://dx.doi.org/10.2478/cl-2021-0006>

CAN CORPUS CONSULTATION COMPENSATE FOR THE LACK OF KNOWLEDGE IN LEGAL TRANSLATION TRAINING?

PATRIZIA GIAMPIERI, MSc.

Università degli Studi di Camerino, School of Law
Via D'Accorso, 16. 62032 Camerino (MC);
Scuola Superiore per Mediatori Linguistici “Carlo Bo”
Via Tommasetti, 6. 00161 Roma, Italy
patrizia.giampieri@unicam.it

ORCID: <http://orcid.org/0000-0003-4443-1593>

Abstract: It is generally assumed that a good knowledge of the legal field is a prerequisite to deliver legal translations. This paper will challenge this assumption by presenting a case study with third-year bachelor's students who participated in a translation project. The students, enrolled in a course in translation practice, were trained in corpus consultation at the beginning of the academic year. Nearly at the end, they translated an extract of a supply contract without being trained in the legal field. They consulted a pre-compiled offline corpus and online bilingual dictionaries. The paper findings highlight that knowledge of the legal field would have certainly helped the students make more informed decisions and avoid some mistranslations. However, the major

shortcomings were actually due to ineffective corpus or dictionary consultation. In particular, formulaic expressions and collocations were neglected. In light of the paper findings, it can be speculated that in translation training, effective corpus consultation may help users deliver high-quality legal translations. It also seemed that thorough knowledge of the legal field is not a prerequisite, at least as far as short texts are concerned.

Keywords: offline corpora; translation training; legal translations; corpus-based translations; DIY corpora.

TRADUZIONI GIURIDICHE: CONSULTARE UN CORPUS PERMETTE DI COMPENSARE LA SCARSA COMPETANZA NEL SETTORE?

Riassunto: Si presuppone generalmente che una certa conoscenza dell'ambito giuridico sia considerata un requisito essenziale nella traduzione di testi giuridici. Il presente articolo mette in discussione il suddetto principio presentando un caso di studio che ha coinvolto studenti iscritti ad un corso di laurea triennale in traduzione. Durante l'anno accademico, i partecipanti hanno ricevuto formazione in linguistica computazionale e nella consultazione di corpora per finalità traduttive. Alla fine dell'a.a., hanno tradotto un estratto di un contratto di fornitura senza tuttavia aver ricevuto formazione in ambito legale. Al fine di svolgere le traduzioni assegnate, si sono avvalsi di un corpus offline pre-compilato e di dizionari online bilingue. L'articolo evidenzia come la conoscenza dell'ambito giuridico avrebbe certamente aiutato gli studenti nella scelta dei traducenti e nell'ovviare a traduzioni inesatte. Tuttavia, l'articolo evidenzia come la maggior parte delle imprecisioni commesse riguardano una consultazione non approfondita del corpus e dei dizionari. Ad esempio, le espressioni formulaiche e collocazioni presenti nel corpus sono state trascurate. Alla luce dei risultati conseguiti, è possibile supporre che nella pratica e formazione traduttiva una consultazione efficace dei corpora specialistici possa aiutare nel processo traduttivo. Se i testi sono brevi, inoltre, si potrebbe ipotizzare che esperienza nel settore legale non sia strettamente necessaria.

Parole: corpora offline; pratica traduttiva; traduzioni giuridiche; traduzioni basate su consultazione di corpora; corpora fai-da-te.

1. Introduction

1.1. Legal discourse

As remarked by many scholars (Tiersma 1999 and 2006; Williams 2011 and 2017; Coulthard and Johnson 2007 and 2010), the language of the law is particularly intricate and excessively verbose for the layperson (Tiersma 1999; Williams 2017). Its archaic constructs and formulaic expressions are challenging and difficult to understand. Some scholars claim, in fact, that legal language is “just ordinary language with a great deal of technical terminology” (Tiersma 2006: 29). For reasons of “clarity and intelligibility” (Williams 2017: 172), the language of the law is hallmarked by technicalities, as “technical accuracy is an essential prerequisite of good justice” (Alcaraz Varó and Hughes 2002: 5). Legal discourse makes an extensive use of the passive voice, archaic terminology, long phrases, syntactic discontinuities, deixis, and it is influenced by Law French and Law Latin (Tiersma 1999 and 2006; Coulthard and Johnsons 2007 and 2010).

Legal translators are, therefore, confronted not only with the intricacies of the legal field, but they also have to tackle the different legal systems related to the source and target language (Šarčević 2000). For this reason, some scholars view legal translators as experts who create a target text which must function in the target culture (Vermeer 1998: 50). Therefore, it is claimed that before approaching legal translations, translators should be acquainted with comparative law (Van Laer 1999). On the other hand, there are scholars who claim that legal translators do not necessarily undertake legal studies, but tend to develop legal expertise by practice and lifelong learning (Prieto Ramos 2020: 29). There are also examples of legal translation projects carried out with academic students without prior training in legal matters (see Vigier Moreno 2016: 104-105). In his research papers, for example, Vigier Moreno (2016 and 2019) remarks the advantages offered by corpus-aided translator training.

1.2. Offline legal corpora

In order to tackle legal translations, several scholars highlight the benefits of consulting online monolingual and multilingual corpora (Pastor and Alcina 2009; Milizia 2010; Vigier Moreno and Sánchez 2017; Giamperi 2018a). Others report that offline corpora help translators make more informed decisions and increase the translation quality (Vigier Moreno 2016 and 2019; Biel 2017; Giamperi 2019a and 2020).

Some offline corpora are also called DIY (do-it-yourself), ad hoc, disposable and on-the-fly corpora (Varantola 2002; Jensen, Moustén and Laursen 2012: 23; Scott 2012: 6). They are a collection of electronic texts compiled for special purposes by the user. They are generally created to translate a specific text and are disposed of after the translation project is over (Zanettin 2002: 242). DIY corpora can be either offline (when consulted offline), or online (when consulted via online platforms).

As anticipated, the use of offline or DIY corpora for legal translations has been discussed by many researchers (Zanettin 2002; Scott 2012; Gallego-Hernández 2015; Vigier Moreno 2016; Vigier and Sánchez 2017; Giamperi 2019a). As a matter of fact, offline and DIY corpora are claimed to be particularly useful in specialised translations. In this respect, scholars report that thanks to corpora, language data are verified systematically, whereas subjectivity and speculation are reduced (Biel 2010: 2). In her research paper, for example, Rodríguez-López (2016) presents the use of corpora in scientific and technical translation training.

Corpora are also insightful because they help discover binomial expressions, collocations and lexical phrases (Vigier Moreno 2016: 100; Vigier and Sánchez 2017: 261). Teubert (2002) claims that multilingual corpora complement dictionaries. This is also highlighted by Andrades Moreno (2013: 2-3), who argues that corpora compensate for the shortcomings of conventional lexicographic tools.

As outlined above, the literature has long remarked the intricacies of legal language (Bhatia 1997; Tiersma 1999; Williams 2011; Giamperi 2016a) which tend to be challenging for the layperson and the translator. In addition, the literature claims that many students in translation studies tend to receive no training in the legal field (Scott 2012: 2; Vigier Moreno 2016: 102). Therefore, when confronted with

legal documents, they are likely to mistranslate or misinterpret words, phrases or terminology (Giampieri 2020). In this respect, corpora can be particularly helpful, especially if used in conjunction with standard language resources (Bowker and Pearson 2002: 14; Jensen, Moustén and Laursen 2012: 21; Giampieri 2018b). For example, dictionaries are often argued to be insufficient or imprecise (Zanettin 2009: 215-217; Scott 2012: 2). In particular, bilingual or multilingual dictionaries cannot always be reliable (Prieto Ramos 2020: 3). This occurs mainly because terms tend to be deprived of their context (Bowker and Pearson 2002:16; De Groot and Van Laer 2006: 65). In a survey on the translation tools used by professionals, Prieto Ramos (2020) highlights that aligned online corpora are consulted by more than 60% of the translators (2020: 4).

As far as the web is concerned, some scholars criticise the common practice of Googling words, because the web is not conceived to retrieve linguistic data and it is very poor in concordancing (Vigier Moreno 2016: 104). Other scholars, instead, claim that Google advanced search and the web as corpus can provide insightful responses in legal translations (Zanettin 2009: 220; Giampieri 2018b). Irrespective of the claim one may wish to endorse, translators need to know how to use the web to retrieve sensible information (Zanettin 2009: 220). For example, the collection of texts for the compilation of a DIY corpus is generally carried out via Google search (Zanettin 2002: 242; Jensen, Moustén and Laursen 2012: 23). Hence, translators, especially the novice ones, should be taught how to search for and collect representative and reliable web data (Zanettin 2009: 220).

Finally, the question of representativeness is highly debated in corpus studies, as it is generally claimed that a corpus needs to be representative of its genre or sub-genre (Biel 2010; McEnery, Xiao and Tono 2010). One might argue that the larger the corpus, the more representative it is. Nonetheless, legal corpora tend to be small (Biel 2010: 4). The reasons are manifold. The first one is confidentiality (Vigier Moreno 2016: 104); one of the major hindrances of building large legal corpora is the private nature of many legal documents (Vigier Moreno and Sánchez 2017: 261). The second reason lies in the fact that legal corpora and legal linguistic studies generally focus on legislation (Biel 2010: 4; see also the work by Bhatia 2010) and EU documents (see Williams and Milizia 2008; Milizia 2010; Biel 2014; and Giampieri 2016b), as no privacy issues are raised on these fields. Therefore, corporate documents and contracts tend to be under-

represented (Scott 2012: 2), despite their high recurrence (Jacometti and Pozzo 2018: 197ff). Another reason why legal corpora are rather small, is the fact that legal discourse is so conservative that it does not need a large corpus to be represented (Bhatia, Langton and Lung 2004: 227). In practice, legal corpora are composed of a few texts which suffice to represent recurrent patterns, collocational use, formulaic expressions and typical phraseology. In his research, Zanettin (2002: 244) reported that students' offline corpora were made of 10-50 texts. Also Giamperi (2019b) carried out a trial lesson with students using a corpus of a dozen texts.

1.3. Corpora in translation training

The literature often distinguishes between two different approaches in using corpora for translation training: corpus use for learning to translate and learning corpus use to translate (Beeby et al. 2009: 1). In the first approach, the lecturer compiles a corpus with preselected tailored data; the corpus will then be consulted and analysed in class. In the second approach, students first compose a corpus, then consult it for translation purposes.

Irrespective of the approach a lecturer might wish to endorse, corpora are claimed to be successfully used in translation training (Monzó 2008; Biel 2010: 12; Zanettin et al 2014; Frankenberg-García 2015; Rodríguez-Inés and Gallego-Hernández 2016). In her research paper, Rodríguez-Inés (2010) highlights how corpora can help develop students' translation competences. In particular, she posits that a systematic use of electronic corpora and other ICT tools are useful in a task-based lesson. Malmkjær (2014) reports the advantages and disadvantages of using corpora in translation training, and Frankenberg-García (2015) describes the benefits and challenges of teaching master's students how to consult corpora for technical translations. Furthermore, Biel (2017) explains how corpora can be used to help students reflect on past translation choices. She also posits that corpora can be consulted in class during the translation process in order to help students develop critical thinking. A similar study is reported by Bowker (1998). In her research paper, she focuses on specialised monolingual native-language corpora and claims that they are an effective translation resource.

2. Purpose of research and research question

In light of the above, this paper explores whether legal translations can be carried out by consulting an offline DIY corpus of *ad hoc* documents. In particular, it challenges the assumption that translators must be trained in the legal field before engaging in legal translations. Furthermore, this paper sheds light on the importance of analysing a corpus effectively in order to retrieve relevant information and increase the translation quality.

In order to do so, a translation project is presented. The project is carried out with 24 bachelor's students enrolled in the third year of a course in translation practice within the same University. All students, hence, followed the same academic translation programme. At the beginning of the academic year, the students (all Italians) attended 6-hour introductory classes on Google advanced search and corpus consultation. During the academic year, they applied Google advanced search techniques in translation assignments covering various topics, such as architecture, tourism and economics. They also had the chance to translate technical and non-technical texts by consulting pre-compiled electronic corpora. In this way, they became acquainted with simple and multiple word search, lemmatisation, concordancing and collocational search. The translation project was carried out at the end of the academic year.

Offline corpus studies are generally carried out for qualitative rather than quantitative work (Corpas-Pastor 2004: 236 quoted in Gallego-Hernández 2015: 376). Therefore, as can be easily inferred, this study is mainly qualitative.

Hopefully, the analysis will shed light on the advantages and disadvantages of using offline DIY corpora in legal translation training and it will provide insightful data on corpus-based legal translations.

3. Methodology

For the purpose of this paper, the students translated an extract of a legal text (a supply agreement) of approximately 130 words from English (their second language) into Italian (their first language). To complete

the task, they consulted online bilingual dictionaries and an offline corpus. The corpus was a single specialised monolingual DIY corpus composed of 48 files (18,541 word types; 297,055 tokens). The corpus was pre-compiled by the lecturer by using BootCaT freeware software (Baroni and Bernardini 2004). BootCaT allows to retrieve documents from the web by simply inputting a set of keywords, namely “seeds”. For the purpose of this project, the BootCaT seeds were the following: *beni mobili*, *contratto di compravendita*, *contratto di fornitura*, *contratto di somministrazione beni*, *internazionale*, *merce*, and *prodotti* (back-translated: movable property, sales and purchase contract, supply contract, contract for the supply of goods, international, merchandise, and products). Not only does BootCaT source and convert documents, but it also builds a corpus automatically. In this way, a corpus of Italian texts related to supply contracts was composed. The students analysed the corpus by means of AntConc freeware software (Anthony 2019). They were allowed to consult the corpus and one or more bilingual dictionaries. It goes without saying that the approach followed in this translation project was the one described by Beeby et al. (2009: 1) as “corpus use for learning to translate”. In this way, corpus analysis skills are instrumental in developing translation competence.

4. Analysis

An analysis of the challenges of the source text and of the students’ (mis)translations will now follow. The source text is reported in Appendix 1. The next paragraphs will shed light on the most challenging terms and phrases.

4.1. Challenges of the source text

The challenging terms and phrases of the source text are reported in Table 1 here below.

Table 1. The challenging phrases and sentences in the source text

In the event of either party failing to meet their contractual obligations under this agreement the other party has the right to terminate the contract at 3 months' notice.
Under this agreement the other party has the right to terminate the contract at 3 months' notice unless such breach of contract is remedied by the defaulting party to the reasonable satisfaction of the non-defaulting party.
If any material breach is committed by either party (...) the non-defaulting party may terminate this agreement.
Any dispute or claim arising out of or in connection with this agreement or its subject matter shall be governed by and construed in accordance with the law of England and Wales.

As can be noted, many phrases are composed of recurrent key words, such as “this agreement”, “terminate” and “breach”. It will be interesting to explore how the students tackled them and what they proposed as translation candidates.

4.2. Translation candidates and corpus analysis

This subsection highlights the translation solutions proposed by the students and it comments on their (mis)translations. In particular, the number and percentages of shortcomings, or of acceptable translations, are presented. Translations are considered acceptable or good when they fulfil the same function in the target language as the original text in the source language (Ordudari 2008). The translation candidates are analysed on the basis of their relevance (i.e., the number of students that

rendered a particular term wrongly).

Appendix 2 reports a detailed description of the translation candidates (either correct or wrong) proposed by the students, together with the number of mistranslations and their frequencies. Appendix 2 also speculates on the reasons why some students probably mistranslated terms or phrases (see the last column).

As far as the noun phrase “the non-defaulting party” is concerned, 16 students out of 24 (67%) proposed some sorts of simplifications, such as *controparte*, *controparte adempiente*, *parte lesa*, and *parte adempiente* (back translations: counterparty, fulfilling counterparty, injured party and fulfilling party). In other instances, they proposed more complicated candidates such as *controparte non inadempiente* (back translation: non-defaulting counterparty). These mistranslations were probably due to the students’ lack of knowledge of the language of the law and in particular of contract law. However, by searching for *parte* (back translation: party) in the corpus, it is possible to obtain more than 800 hits, amongst which *parte non inadempiente* and *parte non ottemperante* (both back translated: non-defaulting party) can be noticed. Therefore, the students could have guessed that a literal translation of “non-defaulting party” was in use. As stated above, they probably did not consider this possibility for lack of knowledge of the language of the law, lack of intuition and/or low self-confidence.

As for the verb phrase “terminate this agreement”, it is evident that the students’ mistranslations (14 students, 58%) were due to lack of knowledge of contract law. As the literature suggests, in fact, this verb phrase gives rise to many misinterpretations in the Italian language (Giamperi 2016a), even purported by multilingual platforms. Therefore, the translation candidates proposed were several, such as *rescindere dal contratto*, *recedere dal presente accordo*, *terminare questo accordo*, *concludere tale accordo*, and *sciogliere il contratto* (back translations: rescind from the contract, withdraw from the present agreement, end this agreement, conclude this agreement, and dissolve the contract). In this case, consulting the corpus would have not been enlightening. However, almost half of the students (10; 42%) were able to rightly translate the phrase *risolvere il contratto*, as they noticed that *risolvere* collocated with *inadempimento del contratto* (back translation: breach of contract) in the corpus.

As for “breach of contract”, 13 students (54%) proposed various translation solutions, such as *violazione del contratto*,

violazione contrattuale, *violazione*, *non ottemperanza del contratto* (back translations: violation of contract, contract violation, violation, non-compliance with the contract). These translation candidates, although partly acceptable, are not precise rendering of “breach of contract”. As the Hoepli and Garzanti dictionaries propose, in fact, the correct translations are *inadempimento* or *inadempienza*. Corpus evidence corroborates that *inadempimento* is the most used (142 hits).

As far as *violazione* is concerned, instead, it collocates with *obbligo* / *obblighi* (13 hits) (back translations: obligation/obligations), or *diritto* / *diritti* (6 hits) (back translations: right/rights). On the contrary, *violazione* (back translation: violation) does not frequently collocate with *contrattuale* (1 hit) (back translation: contract – modifier) or *accordo* (1 hit) (back translation: agreement). Therefore, the imprecise translations proposed by some students were mainly due to inaccurate dictionary consultation and inaccurate corpus analysis.

As regards “this agreement”, 50% of the students (12 out of 24) translated “agreement” literally and proposed *accordo*. Unfortunately, under the Italian civil code, an *accordo* is only a prerequisite of a binding contract (art. 1325 of the Italian civil code). Therefore, “agreement” is too general as it should have been rendered *contratto* (back translation: contract). This could have been guessed by searching for “agreement” in the Hoepli or Garzanti online dictionaries and then it could have been corroborated by corpus evidence. As a matter of fact, the search for *contratto* / *contratti* (back translation: contract OR contracts) in the corpus yielded to more than 3,000 hits; whereas the search for *accord* / *accordi* (back translation: agreement|agreements) to only 290. This misinterpretation was likely to be due to the students’ lack of accuracy in searching for words both on online dictionaries and in the corpus.

The head noun “subject matter” in the sentence “any dispute or claim arising out of or in connection with this agreement or its subject matter” was mistranslated by 12 students (50%). This phrase was very challenging as not many dictionaries list it. For example, only the Garzanti dictionary specifies that its translation is *oggetto* (back translation: object) in the legal field. Half of the students, instead, translated it literally and proposed *materia*, *contenuto*, *soggetto in materia*, *argomento di discussione* (back translations: matter, content, subject in matter, discussion topic). Mistranslating “subject matter” was, hence, partly due to lack of knowledge of contract law, but also to insufficient dictionary search or corpus analysis. For example, by

looking for *qualsiasi controversia* (back translation: any dispute), it is possible to note that the phrase “subject matter” is omitted in the corpus, as in the following sentence: *qualsiasi controversia derivante dal presente contratto o collegata allo stesso* (back translation: any dispute arising out of the present contract or connected with it). One student, in fact, proposed *o conseguente a esso* (back translation: or related to it).

The formulaic expression “any dispute or claim arising out of or in connection with this agreement” was mistranslated by 9 students (38%). Therefore, the majority of them (namely 15; 62%) successfully consulted the corpus and found corresponding formulae such as *qualsiasi controversia o pretesa / reclamo derivante da o connessa con il / relativa al / collegata al presente contratto* (back translations: any dispute or pretension / complaint deriving from or connected with / related to / linked with the present contract). By searching for *controversia* in the corpus, the following formulae come to the fore: *ogni controversia derivante dal presente contratto o comunque ad essa connessa* (back translation: any dispute arising out of the present contract or, in any case, connected to it), and *qualsiasi controversia connessa al, o comunque derivante dal presente contratto* (back translation: any dispute connected to, or in any case arising out of the present contract). Therefore, the reason for the students’ mistranslations, was probably due to ineffective corpus consultation. Those students who did not consult the corpus effectively, in fact, proposed deviant translation solutions, such as *ogni eventuale causa o ricorso del presente contratto*, or *qualsiasi contestazione e richiesta relativa al presente contratto* (back translations: any suit or petition of the present contract, any objection or request related to the present contract).

Also the verb “remedied” in the phrase “such breach of contract is remedied” posed challenges to some students (9; 38%). In particular, this verb was rendered quite fancifully with *offrire compensazione*, *essere di rimedio*, *essere risolta*, *essere risarcito*, and *essere riparata* (back translation: offer compensation, be of remedy, be solved, be awarded damages, be repaired). Also in this case, the students should have consulted dictionaries more effectively and corroborated the translation proposals in the corpus. The Hoepli and Garzanti dictionaries, in fact, suggest a literal translation (*rimediare*), which is present in the corpus and collocates with *inadempimento* (back translation: breach of contract). The majority of students (62%), however, translated “remedied” correctly.

The noun phrase “material breach” was mistranslated (and probably misinterpreted) by 7 students (29%). Apart from the term “breach” which was already commented above, the adjective “material” was rendered without too much critical thinking. Some literal translations were proposed, such as *materiale*, *effettivo* or *fondamentale* (back translations: material, effective and fundamental). It goes without saying that dictionary consultation did not help in this case. As a matter of fact, the candidates listed by several online dictionaries coincided with the terms suggested by the students. However, they did not verify whether the adjectives proposed were used in legal discourse. They could have grasped the right translation of *materiale* by analysing words in context in the corpus. Instead, the students did not take the time to search for the collocations of *inadempimento* (or of the proposed *violazione*) in 1L (first left) position. By searching in the corpus, the adjective *grave* (back translation: severe) would have come to the fore. Strangely enough, most of the students searched and translated the noun phrase “material breach” correctly (71%), despite the higher mistranslation frequencies of the noun phrase “breach of contract” (namely, 54%).

A few students (4; 17%) misinterpreted and mistranslated the phrase “reasonable satisfaction” as they proposed, again, quite fanciful translations, such as *sufficientemente convincente*, *ragionevole opinione*, *ragionevole parere*, *corretta realizzazione* (back translations: sufficiently convincing, reasonable opinion, correct realisation). In this case, a literal translation corroborated by corpus evidence would have sufficed. The students did not apparently rely on their intuitions. This was probably due to their lack of experience or insufficient self-confidence.

Also the phrase “under this agreement” could have been rendered properly by effective corpus consultation. Some students (3; 12%) translated this recurrent formula with infrequent words, such as *come da contratto* and *connesso al presente accordo* (back translations: as per the contract, connected with the present agreement). By searching for *contratto* (back translation: contract) and noting the words in the second and third left position, it was possible to note *ai sensi* before *contratto*, in the phrase *ai sensi del presente contratto* (back translation: under the present contract). Furthermore, the Hoepli dictionary suggests several translation candidates, such as *in virtù di*, *in base a*, *secondo*, *conformemente a*, *ai sensi di*, which could have been corroborated or confuted by corpus analysis.

Three students (12%) mistranslated the phrase “at three months’ notice”: *almeno tre mesi prima, trascorsi tre mesi dalla notifica* (back translations: at least three months before, after three months from the notice). Needless to say, there was actually no legal meaning in such a phrase, and the mistranslations were due to inaccuracy. In this regard, dictionary entries and corpus evidence could have provided insightful answers. If those students had searched for “notice” in the Garzanti or Hoepli dictionaries, they would have noted *preavviso*. By searching for *preavviso* (back translation: notice) in the corpus, they would have found concordances such as *con preavviso di almeno 30 giorni* (back translation: with at least 30 days’ notice). As a matter of fact, it should be pointed out that the majority of the students (88%) translated this phrase correctly.

Finally, only 2 students (8%) did not render the formulaic expression “failing to meet their contractual obligations” properly. In particular, they proposed *fallire nel rispettare i propri obblighi*, and *non rispettare i vincoli contrattuali* (back translation: be unsuccessful in respecting one’s obligations, not respecting the contract bonds). Although these translations might seem adequate, they do not take into consideration the fact that there are fixed formulaic expressions in Italian. In order to find them, the students could have looked for *obblig* contratt** (back translation: oblig* contract*) in the corpus. As a matter of fact, almost all students (22; 92%) translated this formula correctly. For example, some proposed *inadempimento degli obblighi contrattuali* (back translations: non-fulfilment of the contractual obligations).

4.2.1. Reasons for the students’ mistranslations

It is now interesting to speculate on the possible reasons for the students’ shortcomings.

As anticipated, and as reported in Appendix 2, the main reasons were probably due to 1) lack of training or knowledge of contract law; 2) inaccuracy and carelessness; 3) insufficient dictionary search, and 4) insufficient or ineffective corpus analysis (see Table 2).

Table 2 here below summarizes these findings.

Table 2. Reasons for mistranslations

Reasons for mistranslations	Phrases/words involved (and percentage of students' mistranslations)	Number of phrases/words involved
Lack of training or knowledge of contract law	Terminate the agreement (58%), subject matter (50%)	2
Inaccuracy or carelessness	Non-defaulting party (67%), at three months' notice (12%)	2
Insufficient dictionary search	Breach of contract (54%), subject matter (50%), this agreement (50%), remedied (38%), under this agreement (12%)	5
Insufficient or ineffective corpus analysis	Non-defaulting party (67%), breach of contract (54%), this agreement (50%), remedied (38%), any dispute or claim arising out of or in connection with this agreement (38%), material breach (29%), to the reasonable satisfaction (16%), at three months' notice (12%), failing to meet their contractual obligations (8%), under this agreement (12%)	10

The first type of shortcomings (i.e., “lack of training or knowledge of contract law”) concerned words, phrases or head nouns such as “terminate the agreement” and “subject matter” which the students were not aware of. In practice, they probably did not know the legal institutions behind these words and could not grasp their meaning. The second type of shortcomings (i.e., “inaccuracy or carelessness”) revolved around words or expressions which the students could have rendered well in the target language, had they paid more attention to their grammar forms and/or contexts. The third type of shortcomings (i.e., “insufficient dictionary search”) was due to inaccurate dictionary

search. In practice, in those cases, the students could have found the right translation candidates, had they checked dictionary entries more carefully. The fourth type of shortcomings (i.e., “insufficient or ineffective corpus analysis”) was mainly due to the students’ inexperience in corpus consultation. In these cases, they probably consulted the corpus superficially and did not investigate collocations or word occurrences further.

It can be speculated that the lack of training or knowledge of contract law possibly influenced the translation process. Inaccuracy and carelessness, as well, influenced the students in some circumstances; in particular, as far as “non-defaulting party” and “at three months’ notice” are concerned. Insufficient dictionary search, instead, affected the translation of words or phrases such as “subject matter”, “breach of contract”, “this agreement”, “remedied”, and “under this agreement”. These words, in fact, were present in dictionaries but were probably not investigated thoroughly. Finally, it is evident that the majority of the shortcomings were due to insufficient or inaccurate corpus analysis. Many formulae, for example, could have been found by consulting the corpus more accurately (as some students did). Finally, phrases such as “any dispute or claim arising out of or in connection with this agreement”, “to the reasonable satisfaction” and “failing to meet their contractual obligations” could have been grasped by means of corpus analysis.

These are obviously speculations and reflect the views and impressions of the author, who was the students’ lecturer.

5. Findings and discussion

In light of the findings outlined above, it is evident that the students’ mistranslations could have been partly avoided by improving corpus consultation. Therefore, differently from what one could imagine, insufficient training in the legal field was not the main reason for the students’ shortcomings. As Table 2 above reveals, in fact, inaccurate corpus analysis affected the translation of the majority of words/phrases (namely, 10). The second main reason for the students’ shortcomings, was poor dictionary search.

These findings are revealing, as they highlight how training in

corpus analysis is fundamental in order to tackle technical texts. Therefore, it can be speculated that in translation training, deep knowledge of a specific field seems not to be the mainstay of high-quality translations, as long as *ad hoc* corpora are available and users consult standard language resources (such as dictionaries) effectively. These are only assumptions drawn on the basis of this case study. It is self-evident that further research is called for in order to corroborate them. In support of these findings, Prieto Ramos (2020: 29) reports that many translators do not generally have any formal training in legal studies. Hence, they tend to develop legal knowledge during the course of their practice and/or through training.

Research in other sectors could be carried out with a larger number of users, and/or with experienced translators or students at master's level. Furthermore, this translation project only took a 130-word document into consideration, without accounting for the many possible technical terms a translator may encounter in a longer document.

What this case study brought to the fore, was the fact that standard (i.e., non-technical) dictionaries are sometimes of little help. For example, “subject matter” and “material breach” were hard to find, and without proper corpus analysis, no valid candidates could be found.

This translation project also remarked that the translator's self-confidence is sometimes fundamental in order to find appropriate translation solutions. In some cases (e.g., the translation candidates of “remedy” and “reasonable satisfaction”), the students could have relied on their intuition and looked for literal translations. In these cases, they would have found concordances which corroborated their assumptions. Hence, corpus evidence would have “surprised them in unexpected manners” (Zanettin 2001: 184).

6. Conclusions

Twenty-four bachelor's students enrolled in the third year of a course in translation practice took part in a translation project. They translated an extract of a supply contract of approximately 130 words from English into Italian (their mother tongue). To do so, they consulted one or more bilingual dictionaries and analysed a pre-compiled offline corpus. The

corpus was created by using the BootCaT freeware software (Baroni and Bernardini 2004) and it was composed of 48 files, 18,541 word types and 297,055 tokens.

This paper brought to the fore the fact that many students mistranslated phrases (e.g., “breach of contract”, “material breach”) and formulae (e.g., “under this agreement”, “failing to meet their contractual obligations”) mainly because of poor corpus consultation, despite being trained in corpus analysis. The second main reason for the students’ shortcomings was due to insufficient dictionary consultation. Differently from what one may expect, insufficient training in the legal field seemed not to be the main cause of mistranslations. It goes without saying that training in contract law would have helped the students made more informed decisions. However, apparently a large number of shortcomings could have been (and were) tackled by accurate corpus consultation. It is the opinion of the author that corpus-awareness should be raised, in order to help users understand that both intuition and dictionary entries can be worthwhile exploring. It can be speculated that the students’ low self-confidence was another possible hindrance, which did not help them search for possible alternatives.

The findings of this translation project are, hence, useful for the translation training classroom, as they shed light on the importance of corpus analysis rather than on focussing only on sector knowledge. Hence, translation trainers can focus their attention on corpus consultation, analysis of word uses in context and collocational search when preparing students to become professional translators even in a specific field.

This paper does not wish to suggest that technical knowledge is irrelevant or not necessary. On the contrary, it highlights how sector knowledge is relevant, but in the translation training classroom, focus can also be shifted to improving corpus analysis skills.

The limits of this translation project lie in the fact that the document proposed to the students was of only 130 words. A longer text would have posed more challenges and, probably, would have mirrored authentic translation work. Moreover, this project involved bachelor’s students, who might still be too inexperienced to be confronted with complex sector language. Therefore, further research could involve students at master’s level or professionals.

References

- Alcaraz Varó, Enrique, and Brian Hughes. 2002. *Legal Translations Explained*. Manchester, UK / Northampton, MA: St. Jerome Pub.
- Andrades Moreno, Arsenio. 2013. Internet como fuente para la compilación de corpus jurídicos. *Enlaces* 15: 1-15.
- Anthony, Laurence. 2019. *AntConc* (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. <http://www.laurenceanthony.net/software>.
- Baroni, Marco, and Silvia Bernardini. 2004. BootCaT: Bootstrapping corpora and terms from the web. *Proceedings of LREC 2004*. http://sslmit.unibo.it/~baroni/publications/lrec2004/bootcat_lrec_2004.pdf (accessed December, 8 2020).
- Bhatia, Vijay Kumar. 1997. Translating Legal Genres. In *Text Typology and Translation*, ed. Anna Trosborg, 203-214. Amsterdam: John Benjamins.
- Bhatia, Vijay Kumar, Nicola M. Langton, and Jane Lung. 2004. Legal discourse: Opportunities and threats for corpus linguistics. In *Discourse in the Professions. Perspectives from Corpus Linguistics*, eds. Ulla Connor and Thomas A. Upton, 203-231. Amsterdam/Philadelphia: John Benjamins.
- Bhatia, Vijay Kumar. 2010. Drafting Legislative Provisions: Challenges and Opportunities. *The Loophole* 3: 5-15.
- Beeby, Alison, Patricia Rodríguez-Inés, and Pilar Sánchez-Gijón. 2009. Introduction. In *Corpus use and translating. Corpus use for learning to translate and learning corpus use to translate*, eds. Alison Beeby, Patricia Rodríguez-Inés and Pilar Sánchez-Gijón, 1-18. Amsterdam: John Benjamins.
- Biel, Łucja. 2008. Legal terminology in translation practice: dictionaries, googling or discussion forums? *SKASE Journal of Translation and Interpretation* 3: 22-38.
- Biel, Łucja. 2010. Corpus-Based Studies of Legal Language for Translation Purposes: Methodological and Practical Potential. In *Reconceptualizing LSP. Online proceedings of the XVII European LSP Symposium 2009*, eds. Carmen Heine and Jan Engberg, 1-15. Aarhus: Aarhus School of Business, Aarhus University.

- Biel, Łucja. 2014. The textual fit of translated EU law: a corpus-based study of deontic modality. *The Translator* 20 (3): 332-355.
- Biel, Łucja. 2017. Enhancing the communicative dimension of legal translation: comparable corpora in the research-informed classroom. *The Interpreter and Translator Trainer* 11 (4): 316-336.
- Bowker, Lynn. 1998. Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study. *Meta* 43 (4): 631-651.
- Bowker, Lynn, and Jennifer Pearson. 2002. *Working with specialized language*. London and New York: Routledge.
- Corpas-Pastor, Gloria. 2004. Localización de recursos y compilación de corpus vía internet: Aplicaciones para la didáctica de la traducción médica especializada. In *Manual de documentación y terminología para la traducción especializada*, eds. Consuelo Gonzalo-García and Valentín García-Yebra, 223-258. Madrid: Arco.
- Coulthard, Malcolm, and Alison Johnson. 2007. *An Introduction to Forensic Linguistics – Language in Evidence*. London and New York: Routledge.
- Coulthard, Malcolm, and Alison Johnson. 2010. *The Routledge Handbook of Forensic Linguistics*. Abingdon: Routledge.
- De Groot, Gerard-René, and Conrad J. P. Van Laer. 2006. The Dubious Quality of Legal Dictionaries. *International Journal of Legal Information* 34 (1): 65-86.
- Frankenberg-García, Ana. 2015. Training translators to use corpora hands-on: challenges and reactions by a group of 13 students at a UK university. *Corpora* 10 (2): 351-380.
- Gallego-Hernández, Daniel. 2015. The use of corpora as translation resources: A study based on a survey of Spanish professional translators. *Perspectives* 23 (3): 375-391.
- Giamperi, Patrizia. 2016a. A Critical Comparative Analysis of Online Tools for Legal Translations. *The Italian Law Journal* 2 (2): 445-461.
- Giamperi, Patrizia. 2016b. Is the European legal English legalese-free? *Italian Journal of Public Law* 8 (2): 424-440.
- Giamperi, Patrizia. 2018a. Online parallel and comparable corpora for legal translations. *Altre Modernità* 20: 237-252.
- Giamperi, Patrizia. 2018b. The web as corpus and online corpora for legal translations. *Comparative Legilinguistics* 33: 35-55.

- Giampieri, Patrizia. 2019a. *Guida agli strumenti per la traduzione giuridica: coropra offline a supporto del traduttore*. Bologna: Filodiritto.
- Giampieri, Patrizia. 2019b. Manual and automatic corpus compilation: A case study for legal translations. *International Journal of Language Studies* 13 (3): 1-16.
- Giampieri, Patrizia. 2020. An offline corpus for legal translations into Italian: a case study with a land lease agreement. *Altre Modernità* 23 (5): 85-97.
- Hüning, Matthias. 2008. TextSTAT (version 2.9) <http://neon.niederlandistik.fu-berlin.de/en/textstat/>.
- Jacometti, Valentina, and Barbara Pozzo. 2018. *Traduttologia e linguaggio giuridico*. Milan: Wolters Kluwer.
- Jensen, Vigdis, Birthe Moustén, and Anne Lise Laursen. 2012. Electronic Corpora as Translation Tools: A Solution in Practice. *Communication & Language at Work* 1 (1): 21-33.
- Malmkjær Kirsten. 2014. On a pseudo-subversive use of corpora in translator training. In *Corpora in translator education*, eds. Federico Zanettin, Silvia Bernardini and Dominic Stewart, 119-134. New York: Routledge.
- McEnery, Tony, Richard Xiao, and Yukio Tono. 2010. *Corpus-based language studies: an advanced resource book*. London: Routledge.
- Milizia, Denise. 2010. A Linguistic Investigation of the Lisbon Treaty. *Studi sull'integrazione europea* 2: 455-474.
- Monzó, Esther. 2008. Corpus-based activities in legal translator training. *The Interpreter and Translator Trainer* 2 (2): 221-251.
- Ordudari, Mahmoud. 2008. Good Translation: Art, Craft, or Science? *Translation Journal* 12 (1). <https://translationjournal.net/journal/43theory.htm> (accessed December, 8 2020).
- Pastor, Veronica, and Amparo Alcina 2009. Search techniques in corpora for the training of translators. In *Proceedings of the international workshop on natural language processing methods and corpora in translation, lexicography, and language learning*, eds. Iustina Ilisie, Viktor Pekar and Silvia Bernardini, 13-20. Borovets: Incoma.
- Prieto Ramos, Fernando. 2020. The use of resources for legal terminological decision-making: patterns and profile variations

- among institutional translators. *Perspectives*.
<https://doi.org/10.1080/0907676X.2020.1803376> (accessed December, 8 2020).
- Rodríguez-Inés, Patricia. 2010. Electronic Corpora and Other Information and Communication Technology Tools. *The Interpreter and Translator Trainer* 4 (2): 251-282.
- Rodríguez-Inés, Patricia, and Daniel Gallego-Hernández. 2016. Corpus use and learning to translate, almost 20 years on. *Cadernos de Tradução, Florianópolis* 36 (special 1): 9-13.
- Rodríguez-López, Clara Inés. 2016. Using corpora in scientific and technical translation training: resources to identify conventionality and promote creativity. *Cadernos de Tradução* 36 (special): 88-120. <https://doi.org/10.5007/2175-7968.2016v36nesp1p88> (accessed December, 8 2020).
- Šarčević, Susan. 2000. Legal Translation and Translation Theory: A Receiver-oriented Approach. In *La traduction juridique, Histoire, théorie(s) et pratique*, ed. Jean-Claude Gémard, 329-347. Genève: Université de Genève.
<http://www.tradulex.com/en/translators/Legal-Geneva2000>
 (accessed December, 8 2020).
- Scott, Juliette. 2012. Towards Professional Uptake of DIY Electronic Corpora in Legal Genres. *Salford Working Papers in Translation and Interpreting* 1-22.
- Tiersma, Peter. 1999. *Legal language*. Chicago, IL: The University of Chicago Press.
- Tiersma, Peter. 2006. Some Myths About Legal Language. *Law, Culture and the Humanities* 2: 29-50.
- Teubert, Wolfgang. 2002. The role of parallel corpora in translation and multilingual lexicography. In *Lexis in Contrast*, eds. Bengt Attenberg and Sylviane Granger, 189-214. Amsterdam/Philadelphia: John Benjamins.
- Uniform Commercial Code (UCC) Act 174 of 1962 § 440.2106. *Contract, agreement, contract for sale, sale, present sale; definitions of certain terms*.
<http://legislature.mi.gov/doc.aspx?mcl-Act-174-of-1962>
 (accessed December, 8 2020).
- Van Laer, Conrad J. P. 1999. Comparatieve Begrippen voor Juridische Vertalers. *Terminologie et Traduction* 65-75 .

- Varantola, Krista. 2002. Disposable corpora as intelligent tools in translation. *Cadernos de Tradução IX – Tradução e Corpora* 1 (9): 171-189.
- Vermeer, Hans J. 1998. Starting to Unask What Translatology Is About. *Target* 10 (1): 41-68.
- Vigier Moreno, Francisco, and Maria del Mar Sánchez 2017. Using Parallel Corpora to Study the Translation of Legal System-Bound Terms: The Case of Names of English and Spanish Courts. In *Computational and Corpus-Based Phraseology. EUROPHRAS 2017. Lecture Notes in Computer Science*, ed. Mitkov Rusian, 206-273. Cham: Springer.
- Vigier Moreno, Francisco. 2016. Teaching the Use of *ad hoc* Corpora in the Translation of Legal Texts into the Second Language. *Language and Law / Linguagem e Direito* 3 (1): 100-119.
- Vigier Moreno, Francisco. 2019. Corpus-assisted translation of specialised texts into the L2. From the classroom to professional practice. *Trans-kom* 12 (1): 90-106.
- Williams, Christopher. 2011. Legal English and Plain language: an update. *ESP Across Cultures* 8: 139-151.
- Williams Christopher. 2017. Exposing Errors and Removing Errors: Pushing the Boundaries in Legal English. *Altre Modernità* 4: 165-175.
- Williams, Christopher, and Denise Milizia. 2008. How (un)readable is the European Constitution? A comparison of the English version and the Italian version. In *Studi in Onore del Prof. Vincenzo Starace*, ed. Andrea Cannone, 2209-2227. Napoli: Editoriale Scientifica.
- Zanettin, Federico. 2001. Swimming in Words: Corpora, Translation, and Language Learning. In *Learning with corpora*, ed. Guy Aston, 177-197. Houston, TX: Athelstan.
- Zanettin, Federico. 2002. DIY Corpora: The WWW and the Translator. In *Training the Language Services Provider for the New Millennium*, eds. Belinda Maia, Jonathan Haller and Margherita Urlych, 239-248. Porto: Faculdade de Letras, Universidade do Porto.
- Zanettin, Federico. 2009. Corpus-based Translation Activities for Language Learners. *The Interpreter and Translator Trainer* 3 (2): 209-24.
- Zanettin, Federico, Silvia Bernardini, and Dominic Stewart. 2014. *Corpora in translator education*. New York: Routledge.

Online dictionaries

Cambridge: <https://dictionary.cambridge.org>

Garzanti: www.garzantilinguistica.it

Hoepli: <https://dizionari.repubblica.it>

IATE: <https://iate.europa.eu/>

Appendix 1. The source text

Duration of contract

In the event of either party failing to meet their contractual obligations under this agreement the other party has the right to terminate the contract at 3 months' notice unless such breach of contract is remedied by the defaulting party to the reasonable satisfaction of the non-defaulting party. If any material breach is committed by either party which, in the reasonable opinion of the non-defaulting party, cannot be remedied within 7 working days the non-defaulting party may terminate this agreement immediately by way of written notice.

Governing law and jurisdiction

This agreement and any dispute or claim arising out of or in connection with it or its subject matter shall be governed by and construed in accordance with the law of England and Wales.

Appendix 2. The translation candidates proposed by the students

Challenging word or phrases of the source text	Students' correct translations	Students' mistranslations	Number and frequencies of mis-translations	Reasons for mis-translations
The non-defaulting party	<i>Parte non inadempiente</i> [Back-translation: non defaulting party]	<i>Controparte, controparte adempiente, controparte non inadempiente, parte adempiente, parte lesa, parte morosa</i> [Back-translations: counterparty, fulfilling counterparty, non-defaulting counterparty, fulfilling party, injured party, the party in arrears]	16 (67%)	2,4
(If any material breach is committed by either party (...) the non-defaulting party) may terminate	<i>Risolvere il contratto</i> [Back translation: terminate / cancel the contract]	<i>Rescindere il contratto, rescindere dal contratto, recedere dal presente accordo, recedere unilateralmente dal contratto, terminare questo accordo,</i>	14 (58%)	1

this agreement		<p><i>concludere tale accordo, sciogliere il contratto</i></p> <p>[Back translations: rescind the contract, rescind from the contract, withdraw from the present agreement, withdraw unilaterally from the contract, end this agreement, conclude this agreement, dissolve the contract]</p>		
Breach of contract	<p><i>Inadempimento / inadempienza</i></p> <p>[Back translation: non-fulfilment]</p>	<p><i>Violazione del contratto, violazione contrattuale, violazione, non ottemperanza del contratto</i></p> <p>[Back translations: violation of contract, contract violation, violation, non-compliance with the contract]</p>	13 (54%)	3,4
This agreement	<i>Il presente contratto</i>	<i>Il presente accordo, accordo, questo accordo</i>	12 (50%)	3,4

	[Back translation: the present contract]	[Back translations: the present agreement, agreement, this agreement]		
(Arising out of or in connection with its agreement or its) subject matter	<i>Oggetto, o conseguente a esso</i> [Back translations: object, or related to it]	<i>Materia, contenuto, soggetto in materia, argomento di discussione</i> [Back translation: matter, content, subject in matter, discussion topic]	12 (50%)	1,3
Any dispute or claim arising out of or in connection with this agreement	<i>Qualsiasi controversia o pretesa / reclamo derivante da o connessa con il / relativa al / collegata al presente contratto</i> [Back translations: Any dispute or pretension / complaint deriving from or connected with / related to /	<i>In caso di controversia o pretesa risultati dal presente accordo, ogni eventuale causa o ricorso del presente contratto, qualsiasi contestazione e richiesta relativa al presente contratto, qualsiasi controversia o richiesta derivanti dal presente accordo, ogni controversia o reclamo che sorge fuori o in</i>	9 (38%)	4

	linked with the present contract]	<p><i>relazione a questo contratto,</i></p> <p>[Back translations: in case of dispute or claims resulting from the present agreement, any suit or petition of the present contract, any objection or request related to the present contract, any claim or request deriving from the present agreement, any controversy or complaint which arises out of or in relation to this contract]</p>		
(Such breach of contract is) remedied	<p><i>Rimediato</i></p> <p>[Back translation: remedied]</p>	<p><i>Offrire compensazione, essere di rimedio, essere risolta, essere risarcito, essere riparata</i></p> <p>[Back translations: offer compensation, be of remedy, be solved, be awarded damages, be repaired]</p>	9 (38%)	3,4

Material breach	<i>Grave inadempimento / inadempienza</i> [Back translation: severe non-fulfilment]	<i>Violazione effettiva, violazione materiale, violazione sostanziale, grave violazione, inadempimento materiale, violazione fondamentale</i> [Back translations: effective violation, material violation, substantial violation, severe violation, material non-fulfilment, fundamental violation]	7 (29%)	4
(Such breach of contract is remedied ...) to the reasonable satisfaction (of the non-defaulting party)	<i>Ragionevole soddisfazione</i> [Back translation: reasonable satisfaction]	<i>Sufficientemente convincente, ragionevole opinione, ragionevole parere, corretta realizzazione</i> [Back translations: sufficiently convincing, reasonable opinion, correct realisation]	4 (17%)	4
(The contractual	<i>Previste / stabiliti nel</i>	<i>Il presente contratto</i>	3 (12%)	3,4

obligations) under this agreement	<i>contratto, di cui al presente contratto, ai sensi del presente accordo</i> [Back translation: set / established by this contract, under this contract / agreement]	<i>riconosce, come da contratto, connessi al presente accordo</i> [Back translations: the present contract acknowledges, as per the contract, recognised by this agreement]		
At three months' notice	<i>Con un preavviso di tre mesi</i> [Back translation: with a three months' notice]	<i>Almeno tre mesi prima, trascorsi tre mesi dalla notifica</i> [Back translations: at least three months before, after three months from the notice]	3 (12%)	2,4
Failing to meet their contractual obligations	<i>Inadempime nto degli obblighi contrattuali, non rispettare gli obblighi previsti dal contratto, venir meno agli obblighi contrattuali</i>	<i>Fallire nel rispettare i propri obblighi, non rispettare i vincoli contrattuali</i> [Back translations: be unsuccessful in respecting one's obligations, not	2 (8%)	4

	[Back translations: non-fulfilment of the contract obligations, non compliance with the obligations set by the contract, neglect the contract obligations]	respecting the contract bonds]		
--	--	--------------------------------	--	--

Legend (reasons for mistranslations column):

1=lack of training/knowledge of contract law

2=inaccuracy, carelessness

3=insufficient research in dictionaries

4=insufficient or ineffective corpus analysis